ARL

Aviation Research Lab
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

ATTENTION FILTERING IN THE DESIGN OF ELECTRONIC MAP DISPLAYS: A COMPARISON OF COLOR-CODING, INTENSITY CODING, AND DECLUTTERING TECHNIQUES

Michelle Yeh and
Christopher D. Wickens

Technical Report ARL-00-4/FED-LAB-00-2

June 2000

Prepared for

U.S. Army Research Laboratory
Interactive Displays Federated Laboratory

Contract DAAL 01-96-2-0003

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2000		**-**

4. TITLE AND SUBTITLE
ATTENTION FILTERING IN THE DESIGN OF ELECTRONIC MAP DISPLAYS: A COMPARISON OF COLOR-CODING, INTENSITY CODING, AND DECLUTTERING TECHNIQUES

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Army Research Laboratory,Aberdeen Proving Ground,MD,21005

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
In a series of experiments, the use of color-coding, intensity coding, and decluttering were compared order to assess their potential benefits for accessing information from electronic map displays. Participants viewed electronic battlefield maps containing five classes of information discriminable by color, intensity, or in the decluttering condition, displayed or removed entirely by a keypress. Participants were asked questions requiring them to focus on objects within a class (i.e., objects presented at the same color or intensity) or divide their attention between objects in different classes (i.e., objects presented at different colors and intensities). The results suggested that the benefits of color and intensity coding appear to be in segregating the visual field rather than calling attention to the objects presented at a certain color or intensity. Interactivity proved to be a disadvantage; the cost of information retrieval outweighed the benefits of presenting less information on the display or even allowing map users to customize their displays. Potential applications of this research include a cost-benefit analysis for the use of three attentional filtering techniques and presents an attempt to quantitatively measure map complexity.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**31**	

ABSTRACT

In a series of experiments, the use of color-coding, intensity coding, and decluttering were compared order to assess their potential benefits for accessing information from electronic map displays. Participants viewed electronic battlefield maps containing five classes of information discriminable by color, intensity, or in the decluttering condition, displayed or removed entirely by a keypress. Participants were asked questions requiring them to focus on objects within a class (i.e., objects presented at the same color or intensity) or divide their attention between objects in different classes (i.e., objects presented at different colors and intensities). The results suggested that the benefits of color and intensity coding appear to be in segregating the visual field rather than calling attention to the objects presented at a certain color or intensity. Interactivity proved to be a disadvantage; the cost of information retrieval outweighed the benefits of presenting less information on the display or even allowing map users to customize their displays. Potential applications of this research include a cost-benefit analysis for the use of three attentional filtering techniques and presents an attempt to quantitatively measure map complexity.

INTRODUCTION

The presentation of map information on electronic displays is a challenge, due to the amount of information that must be presented constrained by the size of the space in which the information can be displayed. Consider the tasks of the land soldier, who when examining a battlefield map may be potentially searching for friendly and enemy troop locations, assessing the traversability of the terrain, and trying to plan a route. It is impossible to determine in advance what sorts of information need to be depicted; the map user may have a variety of cognitive tasks to be carried out with a specific map, ranging from visual search to location assessment to complex integrative planning. Consequently, the designer of most maps (and displays of all sorts) is confronted by an inevitable tradeoff between information and clutter.

The costs of clutter to visual performance are born most heavily in the task of **visual search** (Teichner and Mocharnuk, 1979), but may also be reflected in **information readout**, when a target item is located, but is difficult to interpret because of the presence of other items in very close (or overlapping) spatial proximity. These costs may be alleviated somewhat by employing attentional filtering techniques. The current study focuses on three such methods: color coding; intensity coding, i.e., foreground-background contrast in which different domains of information are presented at varying levels of brightness; and decluttering, removal of one domain of information which is not relevant to the current task at hand. Color and intensity coding have already been employed on traditional paper maps. For example, the full array of intermediate contour lines are generally presented with less intensity (or less contrast), than the primary lines at 1000 foot intervals. The removal of selected domains of information has been employed in high information density paper approach charts (Hofer, Palen, and Possolo, 1993), but what is also of interest here is the flexibility to *retrieve* those “decluttered” or temporarily “erased” domains of information.

The advantage to each of these attentional filtering techniques can be realized by assessing its benefits on visual search in many map tasks. Visual search often proceeds in serial, in which each item in the “search field” is examined in turn, until the critical target is located. Hence, the more cluttered is the display, the longer, on average, will be required till the target item is found (Teichner and Mocharnuk, 1979). Note that the term “item” on a map is more ambiguous than in typical visual search experiments. For example, is a mountain on a contour map a single item, a number of items each consisting of a defining contour line, or somewhere in between. This issue will be addressed later. The particular benefits of color (and also intensity) coding in this regard, is a result of the nature of preattentive information processing, in which color and intensity discriminations are made relatively early and automatically in the sequence of information processing stages (Treisman, 1988; Treisman and Gelade, 1980). Hence unique colors (or intensity differences) in a search field have been found to produce relatively automatic detection, sometimes known as a “popout” effect (Yantis, 1993). Stated in other terms, search for uniquely colored items appears to proceed more or less in parallel, such that all items of the search field are examined at once and so search is not prolonged as more items of different color are added to the field. Additionally, differing color and intensity facilitates the map user's ability to locate and focus attention on those entities, with reduced distraction from others. In searching menu items or structured “lists”, research by Fisher and his colleagues has demonstrated clearly how “highlighting” a restricted set of items within which a target might be found, can

substantially reduce the search time for a highlighted target to the extent that the highlighting is "valid" (Fisher, Coury, Tengs, and Duffy, 1989; Fisher and Tan, 1989).

The advantages for color versus intensity coding are not clear. Color coding offers better associations (e.g., red = danger) and a greater number of levels associated with meaning (4-6 in operational environments (Wickens and Hollands, 2000) but does not offer as natural an ordering of "importance" as does intensity coding. Colors may be confusable under conditions of highly variable ambient illumination; and because color differences may actually be *more* salient (than intensity levels), their differences may actually hurt the ability to integrate information across differently color coded classes of items, to a greater extent than intensity differences (Wickens and Andre, 1990). Finally, color coding requires slightly more complex and potentially vulnerable technology. In well protected work environments, this might not present a problems, but in field units, electronic failure of some sort becomes more likely, and hence the need to retain component simplicity becomes more important.

The task of map reading is composed of integration or comparison processes which extend beyond that of the search task. When considering the use of color and intensity coding, it is important to note that each of these two techniques may bring with them potential costs in comparison processes, particularly when implemented on electronic maps. According to the *proximity compatibility principle* (Wickens and Carswell, 1995), the greater the extent that sources of information need to be integrated, the more the categories of information should be displayed in close proximity, and if such display proximity is achieved, the less the time and load on working memory that will be necessary to compare and integrate information between items. This reduction may result in less movement by the head and eyes if proximity is achieved by spatial overlap or physical proximity, but may also be realized through "mental" proximity through similarity in the coding of display features, such as the use of a common color or intensity for items that need to be compared. This advantage will be enhanced especially in cluttered displays. Consequently, although the use of unique and discriminable coding may facilitate visual search, this same coding may inhibit information integration, particularly if the objects to be compared are presented in two different colors (Wickens and Andre, 1990). However, Martens and Wickens (1995) reported no such cost to information integration for using different levels of intensity coding, and in fact, found that the benefits of intensity coding for the search task dominated any costs for the integration task. They concluded that differentiating classes of items by intensity coding helped users visually "parse" the complex map into meaningful units and search selectively for items in each domain, before those items needed to be compared.

A third technique to be assessed is that of flexible decluttering, in which domains of map information (e.g., all the interstate highways depicted on a state map) can be removed entirely and recalled when needed. While decluttered displays are clearly easier to read than those that contain clutter (Hofer, Palen, and Possolo, 1993; Mykityshyn, Kuchar, and Hansman, 1994; Schultz, 1986), a full evaluation of the effectiveness of decluttering in operational environments must also consider circumstances in which the information wanted for a particular task happens to be decluttered and therefore "hidden" at the time it is needed, and hence must be recalled by manual intervention. Note that in most studies which have examined decluttering, the displays have been static and intervention was not possible (or required). In the assessment of "flexible" decluttering, before search can commence, the map user must first *decide* whether the domain of

interest – that of which the target item is a member – is displayed; if not, then *retrieving* the relevant information via some interface will be necessary. This task of information retrieval (and the decision to do so) may impose a non-trivial penalty (e.g., a significant increase in the time required to access the information), such that the cost when it is required dominates any benefits of decluttering over a more cluttered display for items that are already present when wanted. Consequently, the opportunities for interactivity may be viewed as more of a burden than an advantage (Oviatt, 1996).

Previous comparisons of the three filtering techniques suggest that the degree to which attentional filtering makes the target information distinctive may play a more critical role in response time than does the density of the displays (Schultz, Nickols, and Curran, 1985). However, it is unclear how map density should be computed. While the benefits of color coding in maps have been clearly demonstrated (Schultz, 1986; Wickens and Yeh, 1997), and while careful quantitative modeling has demonstrated the parallel search (popout effect) for uniquely colored items in a visual search field consisting of spatially localized items (shapes , letters, etc), these two areas have not, apparently been merged. That is, studies have not been done that examined the quantitative benefits of color in turning a time consuming serial search of map features into a parallel search, or at least in restricting the size of the serial search by eliminating from the search field all map items of a different color or intensity. Fisher’s research has examined similar issues in the search task for items in structured lists (Fisher, et al., 1989; Fisher and Tan, 1989), but such research tasks (involving search as only one component) may be less generalizable to comparison in less well structured maps. A further reason for this absence of quantitative modeling of map search is the difficulty in defining what is an “item”. While such definition is relatively straightforward with spatially localized stimuli like letters, geometric symbols, or icons, it is far less so with spatially distributed entities like roads, contour lines and rivers. For example, does a hill made up of four concentric contour lines represent one item or four? Does a curving river make up more “items” than a straight one? Such problems with the heterogeneity of features, items or objects have inhibited efforts to quantitatively define the size of a map search field, and, more generally, metrics for map complexity, although such metrics are well defined in other visual domains, e.g., text pages on instrument panels or menus (Tullis, 1988; Wickens, Vincow, Schopper, and Lincoln, 1997).

The goal of the series of experiments we report here was to assess the effectiveness of these three attentional filtering techniques: color coding, intensity coding, and flexible decluttering. Participants were presented with battlefield maps containing <u>terrain</u> features, <u>linear</u> features consisting of water (rivers and lakes) and roads, and <u>people</u> features consisting of travelling troops and stationary units. Their task was to answer questions which required focused attention on one domain of information, (i.e., items which were presented in the same color or intensity) or divided attention between two or more domains (i.e., compare items which were coded differently or items which might need to be retrieved). Maps appeared either in monochrome, with intensity coding of different feature classes, with color coding of those classes, or with a decluttering option, in which classes could be erased entirely. The display manipulations used in the series of experiments we report are described below in Table 1.

Table 1. Display Manipulations

Condition	**Description**
Monochrome	all information presented in the same color at equal intensities.
Color-coding	levels of information were colored uniquely in three colors: contour lines in green, linear features in blue, and people features in red.
Intensity coding:	each domain of information was presented at varying levels of intensity; contour lines were presented at the lowest intensity whereas the intensity of the linear features and people features varied between conditions. Questions asked required focus on both of the initially highlighted domain - that presented at the highest intensity level - and the lowlighted domains.
highlighting linear features	the map presented linear features at the highest intensity.
highlighting people features	the map presented highlighted the people features – troops and units.
Decluttering	information was presented in the same color and intensity level but domains can be "erased" or displayed at the user's discretion. Although three domains are present, participants were only able to manipulate two of them -- linear features and people features. Contour lines always remained on the display.
partial information present on the initial display	Some questions were asked regarding the decluttered display which could be answered using the information visible on the display at the time the question was posed whereas other questions were asked regarding information in currently hidden subclasses, which needed to be retrieved through keyboard interaction.
decluttering linear features	the map presented showed the terrain and people domains; the linear features were initially hidden.
decluttering people features	the map shown displayed only the terrain and linear features. People information were initially hidden.
full information present on the initial display	All information on the map is presented, and the subject is allowed to declutter the map to better answer the question, by enabling focused attention on the relevant domain with no background clutter.

In Experiment 1, the trade-offs for color-coding, intensity coding, and decluttering were compared on maps of equal complexity. In Experiments 2, we employ the same general paradigm as in Experiment 1 but varied the characteristics of attentional filtering and the amount of clutter (complexity) across a relatively wide range, created by three general complexity levels. As described below, the quantitative measure of map complexity was computed in two different ways. In the latter experiments, we anticipated that increases in map complexity would cause increases in the time to respond to both focused and divided attention questions, and might influence the latter to a greater extent, because twice as many items need to be located in the search field. However we also anticipated that clutter costs for both kinds of questions would be muted when coding was employed, thereby reducing the functional search field (to one class for focused attention questions; to two classes for divided attention questions). Our interest in these studies was in both the nature of the search function for the monochrome displays and in how accurately it could be predicted by the two different "item counts" (using different operational definitions of what constituted an "item"), as well as in the degree of flattening of the complexity search slope when color or intensity coding was added. The goal of Experiment 3 was to assess the usability of flexible decluttering to determine if users would voluntarily declutter electronic maps if such a feature were available, and if doing so would facilitate their search for information.

EXPERIMENT 1

Methods

Participants. Eighteen students at the University of Illinois (10 males and 8 females) served as voluntary participants. All of the participants were paid $5.00 an hour for their participation. All were familiar with topographical map reading.

Displays. Stimuli consist of synthetic battlefield maps, containing five different information categories –terrain, water, roads, troop movement, and unit locations. An example is presented in Figure 1.

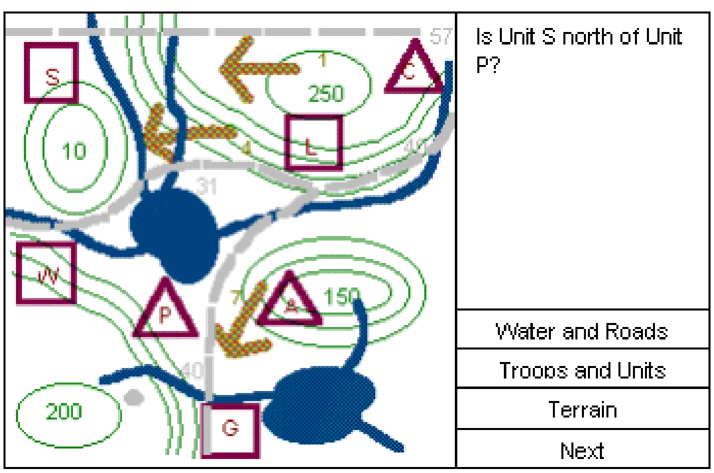

Figure 1. Experiment display.

The maps created can be categorized to define three domains which can be differentiated in terms of both their perceptual and semantic properties:

- Contour lines: terrain
- Linear features: water (rivers and lakes) and roads
- People features: troop movement (arrows) and unit locations (squares and triangles)

The different categories were represented with variations in line texture and shape so as to allow the user to better distinguish the different groups. Contour lines were drawn so that they were only one pixel in line width. Linear features were heavier, drawn with lines six pixels in width; to distinguish between the two sub-domains – water and roads – water features were drawn with thick continuous lines whereas roads were represented with thick dotted lines. Finally "people features", which were located at more restricted areas, were coded by shape – troop movement was presented with arrows whereas unit locations were represented using triangles and squares. Identification codes were then assigned to represent items in each category: contour lines (hills) were represented by three digits, roads were represented by two digits, troop movement was given a single digit, and unit locations were identified by a single letter.

As shown in Table 1, in the color condition, the three information domains were colored uniquely in three colors: contour lines in green, linear features in blue, and people features in red. In the lowlighting conditions, the domains of information were dimmed physically by varying the RGB values. Terrain was always lowlighted, and the degree of this lowlighting was constant (RGB value = 128, 128, 128). The intensity of the linear features and people features were varied between conditions; the highlighted domain was presented in black (RGB value = 0, 0, 0) and the lowlighted domain presented in gray (RGB value = 192, 192, 192). Although the RGB value for terrain suggests that it would appear visible on the display at an intensity between that of the highlighted domain and the lowlighted domain, the one pixel width of the contour lines caused that information domain to appear lighter than the other two information domains. It is important to note however, that when information was lowlighted, it was still well above threshold and therefore easily readable.

The task required participants to search for and answer questions about information presented on the maps either within each category, within each information domain, or between two or more different domains. For example:

Focused attention within a category
Is Unit J in the northeast corner?
Is highway 18 in the east?

Focused attention within a domain
Is Unit P farther south than Troop 4?
Does highway 31 pass north over a lake?

Divided attention between two or more domains
Which unit is located at the highest point on the map?
Is Unit L north of highway 57?

Note that the particular set of questions asked were unique to the map; the order of questions presented on each map was randomized.

Thirty seven maps were developed for the experiment: one map was shown at the start of each condition for practice; the remaining 36 maps were divided into three different groups of twelve so that each subject viewed only a subset the maps. The order of the conditions and order of presentation of the maps was randomized so that each map was presented in each condition.

All participants completed tasks using a Gateway 2000 486/66 personal computer with 16MB RAM and 11” color monitor. As shown in Figure 1, the map display extended the full vertical distance and 75% of the horizontal distance of the monitor; the question was displayed in the remaining 25% to the right of the map. The map remained on the screen while participants answered the question. A legend was included at the bottom right corner; items in the legend represented the different domains present on the map and changed color or intensity to correspond with the current experimental condition. For the decluttered conditions, the legend was color-coded to correspond with the color coding of the response keys.

Procedure. The experiment lasted approximately 1 hour during which the subject was given the instructions for the experiment and then performed the experiment. Since responses to the questions varied, verbal responses, as opposed to keyboard responses, were collected. Participants were instructed to press the space bar once they knew the answer to the question and simultaneously give their verbal response. Response time was calculated from the time the question was displayed to the time participants pressed the space bar. Participants were encouraged to answer quickly but accurately. Participants were given their accuracy measures for the practice trials before starting the experimental trials but were given no indication of performance for the experimental trials. To display the absent information in the decluttering conditions, participants were asked to press the “g” key, which was marked with a blue label, symbolizing water and roads, or the “h” key, which was marked with a red label, symbolizing troops and units

Performance Measures. Accuracy and response time were collected as measures of performance. For the decluttering conditions, data regarding whether the subject displayed the “decluttered” information was also recorded.

Experiment Design. The experiment was a mixed design as shown in Figure 2. Four of the conditions were “fixed” methods such that the user was unable to personally customize the interface: monochrome, color, the intensity coding (highlighting) of linear features, and the intensity coding (highlighting) of people features. Two “flexible” methods were also included: the decluttering of lines and the decluttering of people features. Decluttering enabled the complete removal of the feature in question, or restoration of that feature once it was removed.

Intensity Coding

Fixed	Cell A: Monochrome	Cell B: Color	Cell C: Linear Features	Cell D: People Features
Flexible			Cell E: Linear Features	Cell F: People Features

Decluttering

Figure 2. Experimental design.

Each subject viewed 2 different maps in each of the six conditions, thereby viewing a total of 12 maps. Given that most map-reading tasks require users to look at the same map more than once, each map was used to answer 14 questions consecutively: 5 focused attention questions, 4 questions requiring integration of information within a domain (referred to as <u>integration questions</u>), and 5 divided attention questions requiring divided attention between domains of information. The order of the 14 questions for each map was randomized. To familiarize themselves with the tasks, participants were given 1 map as a practice map at the start of each condition. 6 questions were asked about the map: 2 focused attention, 2 focused attention within a domain, and 2 divided attention between domains.

Results

Two sets of ANOVAs (each conducted on response time and accuracy) were performed: The first was a 4 (display type) X 3 (question type) ANOVA performed on the four fixed displays shown in cells A,B,C, and D of Table 1. The second examined the effects of display interactivity (decluttering), by comparing cells C,D,E, and F.

Fixed Displays: Response Time. The analysis revealed a main effect of question type ($F_{2,30}=99.3$; $p<.01$), indicating that focused attention questions were answered approximately 2 seconds faster (RT=5.5 sec), than the more complex integration and divided attention questions (RT = 7.5 sec). There was no main effect of display mode on response time nor interaction between question type and display.

Fixed Displays: Accuracy. A significant effect of question type on accuracy ($F_{2,30} = 5.73$; $p<.01$) echoed that observed on RT. That is, the two more complex question types involving integration within and between categories showed less accurate performance. While there was no significant main effect of display type on accuracy, $F(3, 45) = 1.95$, $p = 0.14$, separate planned comparisons revealed significantly higher (4%) accuracy in the color coding as compared to the monochrome control condition, $F(1,15) = 5.57$, $p<.05$. The highlighting conditions did not differ significant from each other, $F(1, 15) = 0.02$, $p = 0.90$.

Display Interactivity: Analysis Logic. Analysis of the effect of display decluttering was carried out within the framework of a 2x2 design (see Table 1). One variable in this analysis was the presence of the active decluttering option (Cells E & F in Table 1). Since maps appeared with one information domain or the other initially “hidden”, the keypress could be used to recall the hidden domain if it was needed to answer the focused attention or integrated question. Note that the hidden information domain was *always* needed to answer the divided attention question, since this question category always relied upon information from both domains. The control conditions for this independent variable (Cells C & D in Table 1) were the same highlighting

conditions examined in Analysis 1. Here the given domain of information was "lowlighted" rather than hidden altogether as in the decluttering conditions. The second variable of this 2X2 ANOVA was whether the coding domain of information for a focused attention question was present (or highlighted) or absent (lowlighted) at the time of question presentation.

Display Interactivity: Response Time. The ANOVA on response time revealed a significant main effect of display ($F_{3,45}=4.29$; $p<.01$) indicating a general cost, across all three display types, for the decluttered displays relative to the highlighted displays. Further analysis (Yeh and Wickens, 1997) revealed that two factors amplified this cost. First, the cost was largest for the focused and divided attention questions compared to the integration questions. Second, for both the focused and integrated questions, the cost was substantially larger (around 2 seconds) for questions pertaining information that was initially hidden when the screen appeared, compared to information that was present. A somewhat surprising observation is that there was no **benefit** to focused attention questions pertaining to present information in the decluttering conditions, relative to highlighted information in the intensity coding conditions. That is, totally eliminating irrelevant background information provided no benefits over simply “backgrounding” that information.

Display Interactivity: Accuracy. Generally, error rates paralleled the negative effects of display interactivity on response time, although with reduced magnitude of effect ($F_{6,90}=1.87$; $p<.10$). In fact the presence of interactivity produced a loss in accuracy only for the focused attention questions.

Information Domain. A final observation, characterizing both dependent variables across all display conditions, concerned the generally poorer level of performance on questions regarding people features (troops and units), compared to linear features (roads and rivers). Data analysis revealed increased response times for focused and integrated attention questions concerning people features than linear features, $F(1, 15) = 4.00$, $p = 0.06$ and $F(1, 15) = 30.38$, $p < 0.0001$, respectively. Further details of all of these analyses can be found in Yeh and Wickens (1997).

Discussion

Fixed Displays. The present results were consistent with previous findings that intensity coding, and particularly color coding facilitated performance relative to the baseline control condition, although in the current experiment these effects were not strong (a 2% difference). In particular color coding showed an accuracy advantage over the control condition. It should also be noted that intensity coding conditions were never found to differ significantly from those involving color coding. Somewhat unexpected was the pronounced finding that questions involving linear features (roads and rivers) were considerably easier to answer than those involving people features. This difference may have resulted because there tended to be more people features than linear features on each map, and hence, assuming the search to be a serial one, a longer search was required within a coded class in the former case. The marked differences between the two coding classes, even within the monochrome, equal luminance control condition, suggests that the classes themselves were relatively easy to discriminate on the basis of their spatial properties (i.e., the localized people features versus the more spatially distributed, "curving" linear features). Hence, the coding techniques only offered a redundant

means for segregating what were already somewhat discriminable classes, and as a result the beneficial effects of those coding techniques may have been muted. We also note here that there was no evidence that the existence of discriminable color or intensity codes inhibited the division of attention across the differently coded domains, in contrast to the findings observed by Wickens and Andre (1990), but consistent with those that had been found by Martens and Wickens (1995).

Decluttering. Regarding the decluttering conditions, it is noteworthy but unsurprising that performance on questions when information needed to be retrieved was significantly worse than when that same information was lowlighted; yet performance in the decluttered conditions when the needed information was present, was not better, relative to the condition when the same information was highlighted; that is, the lowlighted irrelevant background information filtered by the attentional filter imposed no penalty on search and readout of the relevant information relative to circumstances when that information was absent entirely. Another way of looking at these results is to assume that any time benefit that may have been gained in the decluttered condition by “hiding” the unwanted information for a particular question, was offset by the time cost of having to decide whether the information present in the question type needed to be retrieved or not.

Thus our results suggest that, in certain circumstances, color coding or intensity coding, utilizing a human attention filter, may be as effective, if not more so, than decluttering, which utilizes a computer filter to parse temporarily relevant from less relevant information. Such an advantage for attentional filters may be amplified under battlefield conditions that are more hostile to keyboard or other manual interactions necessary to call up or erase information. Furthermore we hope to explore whether the faint advantage of color over intensity coding observed here will be amplified or attenuated with more complex maps as we explore in Experiment 2. Certainly the only minimal advantage of color coding over intensity coding found here, would suggest the desirability of the intensity coding technique, when the other costs of color displays, are taken into account.

The results indicated clear benefits of both intensity coding and color coding, relative to the monochrome condition, in which all features were of constant intensity. Such a benefit for distinct colors was observed for divided as well as focused attention questions, replicating the findings for intensity coding of Martens and Wickens (1995), and hence calling into question the applicability of the proximity compatibility principle to circumstances in which close “display proximity” is defined in terms of common color or intensity (rather than space or configurality). We address this issue more fully at the end of this paper. In addition the results indicated a cost for interactive decluttering relative to intensity coding. However the data from this study were silent as to how color affected the search process, because all maps were relatively uniform in their complexity. Unless search fields are varied in terms of the number of items contained thereon, it is impossible to ascertain the added search time consumed as new items are added, and thus distinguish parallel search (no added time) from serial search. In Experiments 2, we employ the same general paradigm but vary the amount of clutter (complexity) across a relatively wide range, created by three general complexity levels in an attempt to model the quantitative affects of clutter. In Experiment 2a, we collect data for monochrome and color-coded displays; in Experiment 2b, we collect data for the intensity-coded displays (linear features highlighted, people features highlighted).

EXPERIMENT 2

Methods

Participants. Thirty-six students at the University of Illinois served as voluntary participants (18 participated in experiment 2a and 18 in experiment 2b). All of the participants were paid $5.00 an hour for their participation. The participants were all student pilots to ensure that all were familiar with topographical map reading.

Displays. A set of 12 maps used in Experiment 1 served as the starting point for the development of the stimuli (e.g., see Figure 1). Features were then removed or added to these maps to create the low clutter (Figure 3a) or high clutter maps (Figure 3b), respectively. Thus, a total of 36 maps were created. Note though that there were only 12 *unique* maps (i.e., defined in terms of the underlying topography). Each map then had three variants of low, medium, and high density. Three additional maps were developed for the practice trials.

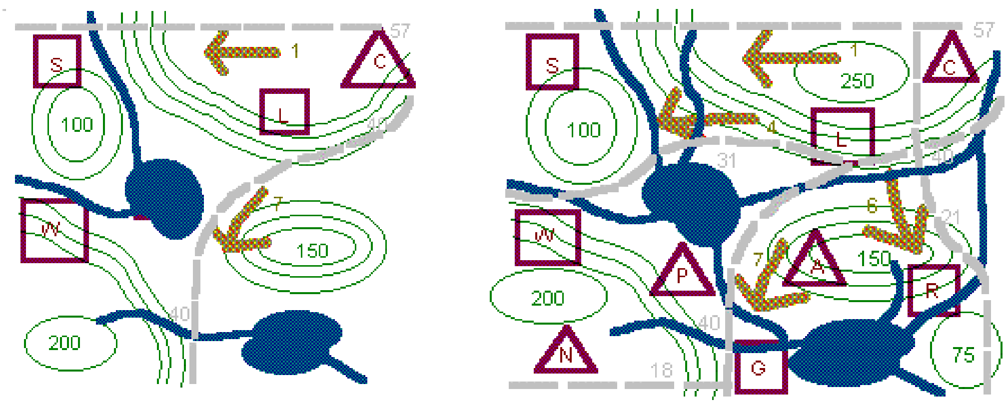

Figure 3. Prototype maps: (a) low clutter, (b) high clutter.

The maps were divided into three different groups of twelve, so that each subject viewed only a subset of the maps. The maps were divided so that participants did not see more than one variant of each of the 12 unique maps; for example, after viewing one unique map in the low clutter condition, that same map (with additional features added) would not be presented for the high clutter trials.

Experiment Design. Each experiment was a 2 [coding: (2a) monochrome vs. color; (2b) intensity coding – people features highlighted vs. linear features highlighted] x 3 (clutter: low, medium, high) within subjects design. However, the general paradigm was that of a mixed design as each subject viewed only a subset of the coding techniques.

Similar to Experiment 1, the maps contained five information categories (terrain, water, roads, troop movement, and unit locations), which were still categorized to define three information domains (terrain, linear features, and people features). However, in the color-coding and intensity-coding conditions, the five levels of information were uniquely coded. When color coding was present, the five levels of information were coded by:

- contour lines = green
- water = blue
- roads = black
- units = red
- troops = brown

When information was intensity coded, five shades of gray were used. The two highest intensity shades were used to depict the highlighted information, and the two lowest intensity shades depicted the lowlighted domains. Terrain information was always coded at the middle intensity level. Note that although the five information categories were coded differently, within each of the two domains, roads were always presented at a higher intensity than water, and units were presented at higher intensities than troops.

In the monochrome condition, all information was presented in black at equal intensities.

Each subject viewed 2 different maps in each of the six conditions, defined by the 3 clutter levels x 2 coding techniques thereby viewing a total of 12 maps. The order of the maps as well as the order of the display conditions were randomized across participants, but each map was presented in each clutter and coding level.

Since the five categories of information were uniquely color-coded (or uniquely intensity-coded), it was impossible to present participants with integrated attention questions (i.e., those questions which required participants to find information about objects across categories but colored identically, e.g., integration between water and roads or troops and units). Consequently, in this study, the class of divided attention questions includes integration within a domain – linear (or people) features, e.g., "does highway 75 cross a river", *and* those requiring integration across domains, e.g., "is Unit B south of highway 44?". Participants were presented with a total of 10 questions: 5 focused attention questions and 5 divided attention questions requiring integration of information between domains.

To familiarize themselves with the tasks, participants were given 1 map as a practice map at the start of each condition. Six questions were asked about the map: 3 focused attention within a domain, and 3 divided attention between domains.

The display presented to participants was identical to that used in Experiment 1 (as shown in Figure 1) with one small exception. Since there were in essence five categories of interest (rather than three general domains as was the case in experiment 1), each category (roads, terrain, troops, units, water – in that order) was given its own place on the legend. That is instead of four buttons to the right of the display, there were now five.

Procedure. The procedure was identical to that used in Experiment 1.

Map Scaling. The complexity of each map was coded in one of two ways: via counting marks and counting objects. The differences in counting affected only the terrain and water features since it was unclear how participants would "group" the features, or even if any grouping would occur. In the mark counting condition, each line was counted as a separate entity. That is, each contour line was counted as one feature and each section of a river was

counted as one feature. For example, if a river ran into a lake and then out of it, the river would be counted as two separate objects. Thus, by this counting rule, Figure 3a shows 10 terrain features, 2 roads, 6 water features, 2 troops, and 4 units.

On the other hand, when counting objects, all terrain features in close proximity which could be used to define one feature (e.g., a hill) were counted as one object. For water features, a river running in and out of a lake was counted as one object. Thus, this counting strategy differs from counting lines in that Figure 3 shows 4 terrain features and 4 water features.

Table 2 lists the mean value and range (standard deviation) of the counts in each of the 2 methods for each of the 3 map clutter categories.

Table 2. Average number of items in each map based on clutter level as determined by the counting marks and counting objects strategies.

	Marks	Object
Low	27.41 (SD = 1.74)	19.13 (SD = 1.04)
Medium	36.34 (SD = 2.44)	25.96 (SD = 2.25)
High	43.01 (SD = 2.63)	32.90 (SD = 1.67)

Results

Overview. Overall, question type had a significant impact on performance for both fixed displays; questions requiring focused attention on one class were answered faster than divided attention questions. Additionally, the results replicated the findings of Teichner and Mocharnuk (1979), who found that increasing the number of items on a display increase the amount of time to locate an item. Accuracy decreased with increasing levels of information, but this effect was somewhat mediated by the use of color-coding. The results are described in greater detail for each type of display and question type.

Response Time. A 3 (clutter: low, medium, and high) x 4 (coding: monochrome, color, intensity – linear features highlighted and people features highlighted) between subjects ANOVA was conducted on the data for both time and accuracy. Figure 4 shows the results for response time.

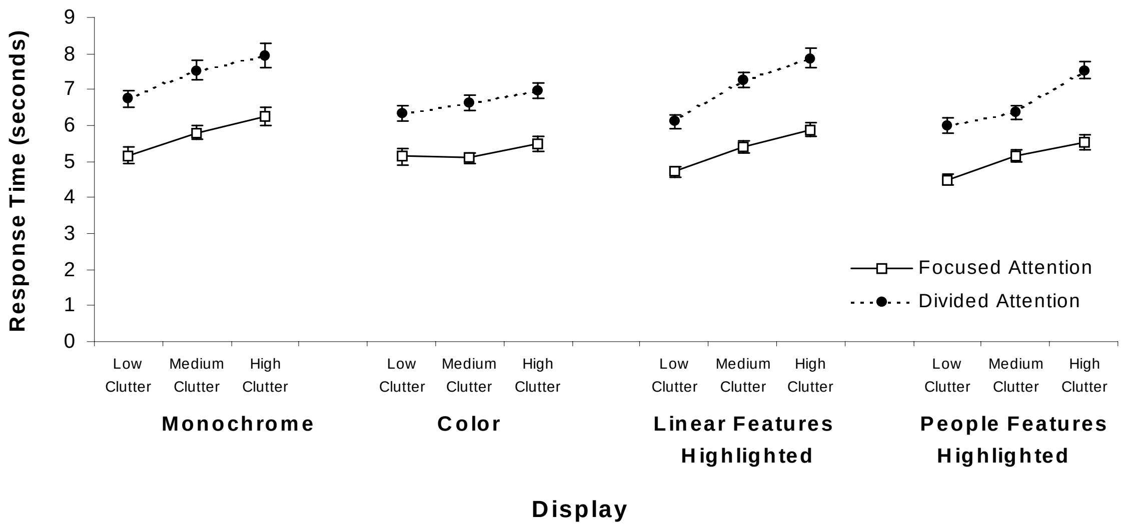

Figure 4. Experiment 2: Response time.

Not surprisingly, as Figure 4 shows, focused attention questions were answered faster than divided attention questions, $F(1, 408) = 108.65$, $p < 0.01$, and increasing the amount of information on the display increased participants' response times, $F(2, 408) = 17.76$, $p < 0.01$. Interestingly, the slightly non-linear trend of the curve for the monochrome condition parallels both "count" variables of map clutter in revealing that the difference between low and medium complexity was greater than that between medium and high (see Table 2).

The data revealed a significant effect of display coding, $F(3, 408) = 4.33$, $p < 0.01$. Tukey's honestly significant difference (HSD) tests, used to compare response time between display conditions, revealed that color coding and intensity coding when people features were highlighted significantly facilitated performance, $p < 0.05$, relative to the monochrome condition.

We were interested not only in how the four display conditions influenced performance relative to each other (as was discussed above) but also in how each of the display conditions modulated the effects of clutter. In order to investigate this latter effect, we analyzed the data for each display condition individually, using a 3 (clutter: low, medium, and high) x 2 (question type: focused attention, divided attention) within subjects ANOVA. The analyses revealed significant effects of clutter on response time when participants were using the monochrome (1.3 s clutter cost, $F(2, 28) = 6.31$, $p < 0.01$), and intensity coded displays [linear features highlighted: 1.4s clutter cost, $F(2, 32) = 13.80$, $p < 0.01$; people features highlighted: 1.3s clutter cost, $F(2, 32) = 13.11$, $p < 0.01$], but not when color coding was implemented, $F(2, 28) = 2.24$, $p = .14$. As the figure suggests, the effects of question type were additive with both clutter and coding.

Accuracy. The accuracy results are presented in Figure 5.

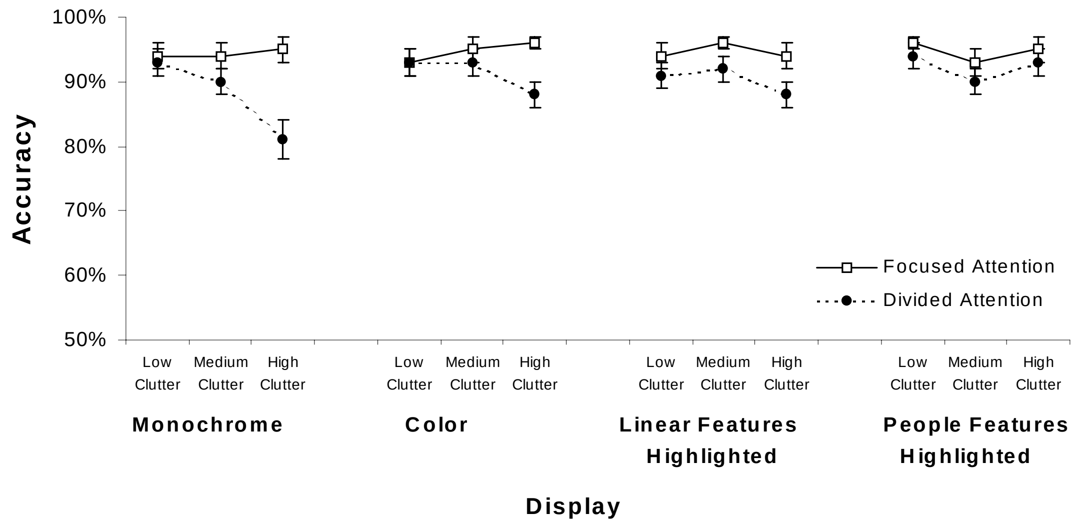

Figure 5. Experiment 2: Accuracy.

As Figure 5 shows, there was a significant effect of question types; the divided attention questions were answered less accurately than those requiring focused attention, $F(1, 408) = 17.68$, $p < 0.01$. There was no overall effect of clutter, $F(2, 408) = 2.11$, $p = 0.12$, nor an interaction between clutter and question type, $F(2, 408) = 2.01$, $p = 0.13$, nor, in contrast to the response time data, was there an effect of display coding, $F(3, 408) = 1.91$, $p = 0.13$. However, the strong visual pattern presented in Figure 5 suggests that the effects of increasing clutter had more impact on the accuracy of answering divided attention questions [$F(2, 204) = 5.79$, $p < 0.01$] than focused attention questions [$F(2, 204) = 0.70$, $p = 0.50$]. The detrimental effects of increasing clutter on answering divided attention questions was pronounced with the monochrome display (a marginally significant 12% loss in accuracy, $F(2, 28) = 2.69$, $p = 0.09$) but nearly absent when display coding was used [a small, nonsignificant 6% loss in accuracy with color coding, $F(2, 28) = 0.52$, $p = 0.59$, and nearly absent with intensity coding (linear features highlighted, $F(2, 32) = 1.92$, $p = 0.16$; people features highlighted, $F(2, 32) = 0.24$, $p = 0.79$)].

Neither the interaction between clutter and coding, $F(6, 408) = 0.26$, $p = 0.85$, nor question type by coding, $F(3, 408) = 1.36$, $p = 0.23$, were significant.

Quantitative Analysis. Following the tradition of previous researchers in visual search, the data were also analyzed via regression models, examining response time regressed on the quantitative measures of visual clutter as assessed by each of the two methods. One set of regression analyses was based upon the average data over participants. These were essentially performed on each of the 4 triads of data points shown in Figure 4, except that the relative spacing of the 3 clutter levels along the x-axis was dictated by the mean calculated complexity level across all 12 maps within each clutter category. As shown in Table 3, because of the procedures used to generate the clutter levels, these values did not quite produce equal intervals between ‘low-medium” and “medium-high”.

Table 3. Regression slopes, intercepts, and R^2 for the four display types.

		Line Count			Object Count		
Averages		Slope	R^2	Intercept	Slope	R^2	Intercept
Monochrome	Focused Attention	0.06	1.0	3.60	0.07	1.0	3.94
	Divided Attention	0.07	0.99	4.75	0.09	0.97	5.19
Color	Focused Attention	0.01	0.23	4.86	0.02	0.30	4.88
	Divided Attention	0.05	0.95	5.04	.05	0.98	5.28
Linear Features Highlighted	Focused Attention	0.07	1.0	2.98	0.08	1.0	3.35
	Divided Attention	0.11	0.98	3.19	0.12	0.95	3.85
People Features Highlighted	Focused Attention	0.07	1.0	2.61	0.08	1.0	3.01
	Divided Attention	0.09	0.81	3.58	0.10	0.86	3.98

The regression slopes, intercepts, and R^2 for each of the four triads, with each complexity coding method are shown in Table 3.

The data shown in Table 3 reveal little difference in the analysis when the regression is performed by counting the individual marks or defined by the number of objects in the display. One may interpret these data as suggesting that high slopes and R^2 values indicate the relevance of the serial search model, i.e.,

$$RT = (IN)/2 \text{ when the target is present, and}$$
$$RT = IN \text{ when the target is absent,}$$

with RT = response time, I = inspection time per item, and N = number of items.

Assuming a serial self-terminating search model and given that in all questions the relevant item was present (i.e., “yes” trials), the search time per item can be estimated by doubling the slope parameters for the focused attention questions. The data then suggest that the serial search model is generally appropriate for describing response time for all the focused attention data but those in the color coding condition, where R^2 drops from around 1.0 to 0.23. The use of highlighting, in contrast, did not appear to filter the search field relative to the monochrome display.

Discussion

The results of the current experiment emphasize the detrimental effects of clutter. Response time to focused and divided attention questions increased as the amount of information presented increased, thus reflecting the amount of time needed to locate the objects in the display. Accuracy decreased as clutter increased, but it is important to consider the interaction between clutter and question type, such that accuracy on the easier focused attention questions was little affected by the addition of information on the display but accuracy on divided attention questions decreased significantly. In other words, clutter imposes a cost, impairing the accuracy of performance when multiple sources of information need to be located and integrated; these are situations in which the information access cost for retrieving information in cluttered displays increases with the added burden of integrating information (Wickens and Carswell, 1995; Schons and Wickens, 1993; Vincow and Wickens, 1993).

The results also point to the benefits of color-coding and intensity coding. The use of color shortened response time and increased accuracy relative to the monochrome display; thus, the results do not show a speed-accuracy trade-off. Additionally, the use of color reduced the negative effects of increasing clutter relative to performance with the monochrome display, as shown in Figure 5, in which the divided attention decrement in accuracy as the level of clutter increases is less in the color-coded condition than in the monochrome condition. The success of intensity coding was less dramatic, but varying the intensity of the information domains buffered the effects of clutter on accuracy, if not on response time. In fact, an examination of the data for divided attention questions in Figure 5 (the bottom lines) reveals that the effect of clutter is essentially eliminated with intensity coding.

The data also speak to the applicability of serial exhaustive search models, developed with simpler stimulus material (Fisher et al., 1989; Neisser, Novick, and Lazar, 1964), to the more complex map material used here. Were these models to fully predict performance in the paradigm we employed, the data should show:

(a) a linear effect of clutter on response time,

(b) a steeper clutter effect on response time for divided than for focused attention questions, since the former would require that approximately twice as many items be located,

(c) a greater benefit for focused than divided attention questions, achieved by color or intensity coding of separate domains of information, given that this filtering would decrease the search field, and

(d) the benefit to coding for focused attention questions described in (c) should be amplified as clutter increased.

Our data are only partially consistent with these predictions. For example, the data do show the clear and linear effect of clutter on response time, as predicted by (a), but do not reveal the enhanced benefit of coding on focused (relative to divided) attention questions as predicted by (c). In fact, this effect is reversed; coding is more beneficial for divided than focused attention questions, and this effect is manifest in accuracy and not in response time and is enhanced at the

highest clutter levels. Finally, the data reveal benefits of color coding in buffering the effects of clutter on response time, but, as shown in Figure 4 (for color coding), this benefit is realized equally for divided and focused attention questions in contrast to prediction (d).

The specific process model that underlies these effects remains somewhat unclear, along with why many of these effects would be manifest in accuracy, rather than response time. It is possible that some aspects of processing (or search) may be carried out in parallel, so that search time is not differentially prolonged by higher clutter in divided relative to focused attention questions. Nevertheless, what is most critical about these data are (1) that a linear search model can be reasonably well applied to map data, (2) that "items" can be easily determined by objects, as by separate marks, and (3) that the coding of differential classes can facilitate performance, sometimes by buffering the effects of increasing clutter (as shown in Figures 4 and 5), and sometimes by "main" effects (i.e., intercepts) that have other benefits on information processing speed and accuracy (as presented in Table 3).

It is important to note that the addition of color and intensity coding did not impair performance on divided attention tasks, i.e., when two differently coded domains of information needed to be integrated. This result is opposite that found by Andre and Wickens (1989) with color coding but similar to those reported in Experiment 1 and by Martens and Wickens (1995) in an examination of intensity coding, in which response time to questions decreased as the intensity differences between domains increased. Martens and Wickens attributed their results to the nature of their task, similar to the task used in the current experiment which emphasized a visual search component to a greater extent than the task employed by Andre and Wickens (1989). That is, for both Martens and Wickens (1995) and the current experiment, the divided attention task was composed of two (or more) focused attention search tasks conducted more or less in series across cluttered domains, whose performance was facilitated by the color differences between the domains, followed by an integration task. Thus, the use of color helped participants focus attention on objects within each domain in turn, and once found, those objects could be compared and integrated. In contrast, the differently coded (colored) indicators used by Andre and Wickens were positioned at the same consistent location in relatively "clutter free" fields, so that no search was required. It was this search process, in Martens and Wickens study and in the current data, that was facilitated by the color distinctions, allowing the "search field" to be pared down.

An alternative solution to the problem of map clutter may be to employ a combination of intensity coding and decluttering, such that the user could highlight (or lowlight) domains of information as it became necessary. One advantage for this technique is the ability to adaptively customize the display, removing layers of information which may not be critical to the given task. An issue that has yet to be examined in great detail is the perceived utility (or cost) of interactivity, e.g., if participants are presented with a highly cluttered display and are given the option to adaptively filter different domains of information. Additionally, the use of intensity coding can be broken down into whether the relevant domain for focused attention was highlighted or lowlighted. Note that this adaptive decluttering is different from the decluttering conditions studied in the previous experiments. For example, in Experiment 1, when the display was decluttered, one domain (either linear features or people features) was initially removed from the display and subjects could only manipulate the presence or absence of that one domain. In Experiment 3, subjects are presented with the ability to adaptively filter either or both domains

of information, and thus for focused attention questions, participants will have the choice as to whether or not they want to remove the irrelevant domain.

EXPERIMENT 3

Methods

Participants. Six students at the University of Illinois served as voluntary participants. All of the participants were paid $6.00 an hour for their participation.

Displays. Stimuli were the 12 high clutter maps used in Experiment 2. Two attentional filtering methods were presented (*highlighting* and *decluttering*), and the implementation of these two coding techniques was such that participants were able to *interact* with an attentionally filtered display or *adapt* the display using one of the two attentional filtering techniques. In the *interactive* conditions, participants were presented at the beginning of each trial with a map with either the linear features highlighted and people features lowlighted (or linear features present and people features decluttered) or vice versa (i.e., people features highlighted and linear features lowlighted or correspondingly, people features present and linear features decluttered). Thus, the interactive conditions involved two versions, decluttering and interactive highlighting. In the decluttering condition, the trial started with one feature class (i.e., linear or people) displayed and the other absent. Participants could display the other class if needed, e.g., in order to answer a divided attention question or a question regarding the focus on the missing class. In the interactive highlighting condition, the feature class was not initially absent (as in decluttering), but was dimmed or "lowlighted" and could then be intensified through the keyboard interaction. In fact, in both interactive decluttering (and highlighting), participants could choose to remove or restore (or lowlight or highlight) either database as many times as they wished.

In the *adaptive* conditions, participants were presented initially with a monochrome map displaying all three information domains (linear features, people features, and terrain at equal intensity) and were given the ability to adaptively declutter (remove entirely) or adaptively lowlight the domains, as they would in the interactive decluttering and highlighting conditions. Thus, in these adaptive filtering conditions, all the questions could be answered without interaction with the display. Note that it was possible to add or highlight both domains so that participants viewed a monochrome map; conversely, it was also possible to remove or lowlight both domains of information so that participants would be left with either a blank display or a lowlighted map.

Each subject viewed 2 different maps in each of the six display conditions. The order of the maps as well as the order of the display conditions were randomized across participants, but each map was presented in each coding level. Participants pressed the "r" (coded with a blue label) and "p" (coded with a red label) keys, to manipulate the intensity (highlighting) or presence (decluttering) of linear features (rivers and roads) and the people features (troops and units), respectively.

The questions presented to participants and the methodology were the same as that employed in Experiment 2.

The data are examined using the framework presented in Figure 6, which presents the initial display states for subjects in the decluttering (Figure 6a) and highlighting (Figure 6b) conditions, and depicts how these display states support the task to be performed. The lines represent the presence of linear features and the circles the presence of people features.

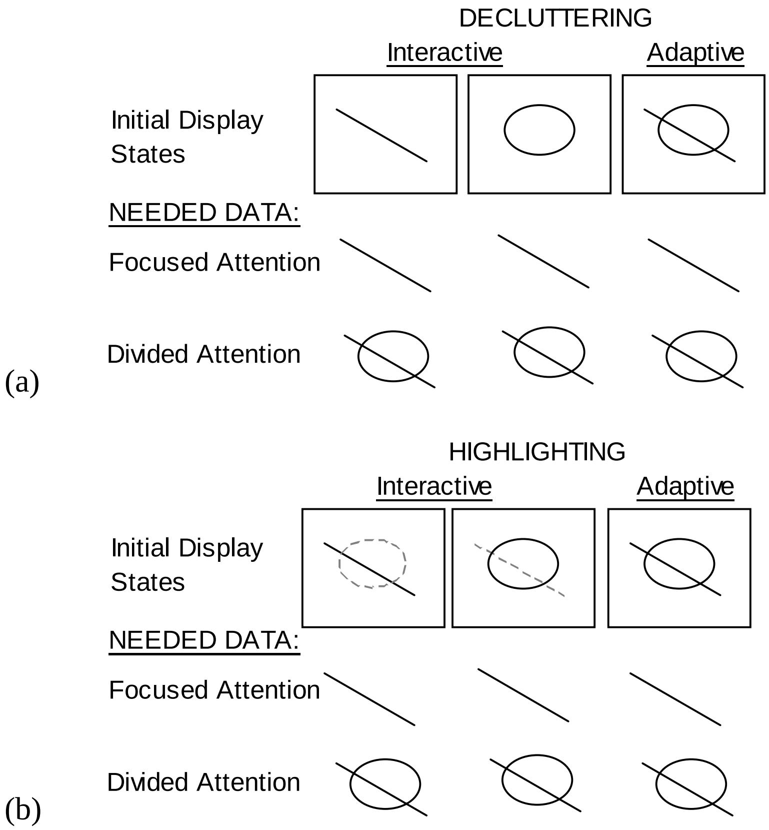

Figure 6. Interactive and adaptive decluttering and highlighting conditions as implemented in Experiment 3. The lines represent linear features and the circles people features.

In the decluttering conditions, participants viewed a map displayed in one of three initial states (as shown in the top row of Figure 6a): in the *interactive* conditions, either the linear features were present or people features present (but not both), and in the *adaptive* decluttering condition, both domains were present. With these displays, participants were required to answer focused attention questions, requiring data from linear features (as represented by the second row of Figure 6a) or people features, or answer divided attention questions, requiring information from both domains (as represented in the third row of Figure 6a). Note that in answering the focused attention question (the second column in Figure 6a displays the needed information – in this example, the linear features), the needed information could either be (1) present (as in the first column), (2) absent (as in the second column) and hence need to be retrieved, or (3) present but in a “cluttered” display (as in the third column).

The framework discussed above for the decluttering condition can be applied to examining performance with the highlighting conditions. In this display condition, the map was presented in the *interactive* decluttering condition, with either the linear features highlighted or

people features highlighted (but not both), and in the *adaptive* highlighting condition, both domains were highlighted. The data required to answer focused attention questions (as presented in the second row of Figure 6b) required attending to the highlighted domain (as in the first column), the lowlighted domain (as in the second column), or filtering the display (as in the third column). Finally, participants were also required to answer divided attention questions in which information from both domains needed to be integrated (as represented in the third row of Figure 6b), and this could occur across two intensity levels (in the first two columns). In comparing highlighting with decluttering, it should be noted that the former involves more discretionary behavior. That is, subjects confronted with the display question situation in the middle column (a focused attention question on a feature that was initially lowlighted) have the option of answering it without any interaction. Participants confronting the same situation in the interactive decluttering condition (the middle column of Figure 6a) do not.

Results

We wanted to assess (1) the frequency of use for the interactivity function, (2) the benefits (or costs) or interactivity as shown be response time and accuracy, and (3) the time cost when the information that was required to perform the task was not present – or lowlighted.

Table 4 presents the data for the frequency with which participants utilized the interactive and adaptive features as a function of question type (focused or divided), and for focused attention questions, whether the needed information was present or absent. It is important to note that in examining the data, participants in the adaptive coding conditions (the two right columns in Table 4) were presented with *all* the information necessary to answer the question, but this information was presented initially without the use of any attentional filtering. Thus, for these two display conditions, the data speak to the utility of the different display coding techniques, as such filtering was not necessary nor required for answering the questions.

Table 4. Frequency of use.

	Interactive		**Adaptive**	
	Highlighting/ Lowlighting	**Decluttering**	**Lowlighting/ Highlighting**	**Decluttering**
Focused attention: Needed present (or highlighted)	(A) 8%	(B) 5%	(C) 30%	(D) 20%
Focused attention: Needed absent (or lowlighted)	(E) 27%	(F) 100%	(G) --------	(H) --------
Divided attention	(I) 18%	(J) 65%	(K) 32%	(L) 14%

Two separate data analyses, depending on question type, were conducted on the data shown in Table 4. As shown in Figure 6, the ability to answer focused attention questions could be separated into two cases: whether the needed information was initially present (or highlighted) or initially absent (or lowlighted). The ability to answer divided attention questions, however, required integrating information across two domains; consequently, one domain of information always needed to retrieved. That is there was no divided attention counterpart to the focused attention condition in which the information was already present.

The data for the focused attention questions were analyzed using a 4 (display) x 2 (presence vs. absence of the needed domain) between subjects ANOVA. Participants interacted most with the display to retrieve information (i.e., in the interactive decluttering condition when the needed information was not initially present); Tukey's HSD pairwise comparisons revealed no other display comparisons to be significant for the focused attention data. It is important, however, to note that participants <u>did</u> use the interactive tools about 25% of the time, when it was not necessary for them to do so (cells C, D, and E in Table 4).

For the divided attention questions, a between subjects ANOVA revealed a significant effect of display, $F(3, 20) = 7.12$, $p < 0.001$. Tukey's HSD revealed that this difference was primarily attributable to the frequency with which subjects added information in the interactive decluttering condition, $p < 0.05$, as of course was required on 50% of the trials. There were no differences in frequency of use between interactive highlighting and adaptive decluttering or lowlighting conditions. However it is important again to note that adaptive highlighting/lowlighting was employed approximately one third of the time (cell K).

The frequency of use data for the interactivity feature only tells half the story, revealing nothing about the potential benefits or costs of interactivity. In order to assess the effectiveness of the attentional filtering techniques employed, the accuracy and response time effects for the adaptive and interactive display coding conditions were compared to a baseline condition, i.e., when interactivity was not available in the monochrome and highlighted displays. No such baseline was included in Experiment 3, but was collected previously in Experiment 2 (the fixed highlighting and monochrome display conditions, viewed in high clutter). For focused attention questions, we were able to compare performance as a function of information presence or absence (or highlighted vs. lowlighted); however, for the divided attention questions, this comparison did not make sense as one domain of information always needed to be retrieved.

The accuracy data revealed no effect of display on answering focused attention questions when the needed information was present, $F(3, 44) = 0.45$, $p = 0.72$, nor when the needed information needed to be retrieved, $F(2, 28) = 0.70$, $p = 0.51$, nor was there an effect of display on the accuracy of answering divided attention questions, $F(3, 75) = 1.92$, $p = 0.13$.

The response time data speak more to the trade-offs of interactivity and adaptivity, i.e., the time costs of using the interactive techniques versus the time benefits of searching a display made less cluttered by those techniques.. The response times for each display condition used in Experiment 3 are presented in Table 5a; the data for the baseline conditions (the fixed monochrome and fixed highlighting displays from Experiment 2) are presented in Table 5b.

Table 5. Mean response time and standard error (in parenthesis). (a) Interactive and adaptive highlighting and decluttering conditions, (b) baseline data: fixed monochrome and fixed highlighting.

(5a)	Interactive		Adaptive	
	Highlighting/ Lowlighting	**Decluttering**	**Lowlighting/ Highlighting**	**Decluttering**
Focused attention: Needed present (or highlighted)	7.99 (0.47)	5.88 (0.29)	6.10 (0.36)	6.94 (0.47)
Focused attention: Needed absent (or lowlighted)	7.37 (0.55)	9.90 (0.69)	--------	--------
Divided attention	10.37 (0.49)	9.84 (0.37)	9.26 (0.41)	9.24 (0.63)

(5b)	**Monochrome**	**Fixed Highlighting**
Focused attention: Needed present (or highlighted)	5.86 (0.21)	5.31 (0.15)
Focused attention: Needed absent (or lowlighted)	--------	5.89 (0.21)
Divided attention	7.40 (0.29)	7.47 (0.17)

The data generally show that allowing one to interact with the display (i.e., with interactive or adaptive highlighting or decluttering) imposes costs, reflected in increased response times relative to the fixed displays (approximately 2s for focused attention questions, $F(1, 44) = 16.52$, $p < 0.001$, and 2s for divided attention questions, $F(1, 44) = 39.74$, $p < 0.001$). In order to investigate the nature of this effect, we compare response times for answering the focused attention questions for the four display conditions (presented in Table 5a) to the two baseline conditions (presented in Table 5b) as a function of whether the information needed was present versus when it was absent (and information needed to be retrieved). For the former (information present), an examination of the data for the interactive highlighting condition (7.99s) relative to the fixed displays (in Table 5b) revealed a cost of interactivity – a 2.7s cost relative to the fixed highlighted display and a 2.13 cost relative to the fixed monochrome display; Tukey's HSD pair-wise comparisons revealed these differences to be significant, $p < 0.05$). No other differences were significant. In particular, the interactive decluttering condition, in which only information relevant to the question domain was presented, supported performance (5.88s)

that was no faster than the control condition in which irrelevant information was either present at the same intensity (5.86s) or reduced intensity (5.31s).

When needed information was absent, the data revealed a significant time cost attributable to information retrieval. Participants using the interactive decluttering condition (i.e., the only display condition in which the needed information was initially hidden) required significantly more time to answer the question relative to the adaptive (9.90s – 6.94s = 3s cost) and fixed displays (a 9.90s – 5.86s = 4s cost); Tukey's HSD revealed this difference to be significant, $p < 0.05$. There was no difference in the time required to answer questions with the interactive highlighting display (7.37s) relative to the adaptive (6.10s) nor fixed (5.89s) displays.

For divided attention questions, the response time data revealed a significant cost of interactivity, $F(5, 54) = 4.86$, $p < 0.001$, attributable mainly to the time cost for interactive highlighting relative to the fixed displays (Tukey's HSD pairwise comparison, $p < 0.05$).

Thus, the data reveal few benefits for interactivity. Although there was no cost when the data was removed (or hidden) at the onset of a trial relative to the monochrome condition when the information needed was already present (for focused attention questions, 5.88s vs. 5.86s, respectively, as shown in Table 5), there was also no benefit for decluttering. Additionally, the effect of the total erasure of clutter (i.e., with interactive decluttering) must be balanced against the substantial cost of retrieving the hidden information when it was the desired focus of attention [9.90s-5.86s = 4.04s]. The use of interactive highlighting resulted in similar costs in information "retrieval" relative to the fixed highlighted display: 7.99s – 5.31s = 2.6s for focused attention questions and 10.37s – 7.47s = 2.9s for divided attention), though surprisingly, this occurred when the needed information was already highlighted (as shown in the first row of Tables 5a and 5b). In contrast, there was no penalty, and even a slight benefit for answering questions when fixed lowlighting was displayed.

Discussion

The data reinforces the success of fixed highlighting for filtering information, allowing a means for the display designer to overlay information without hindering the focus of attention on one information domain or another. The ability to remove layers of information resulted in performance as good as the baseline conditions *when the needed information was present*. However, the data revealed that interactivity resulted in a significant time cost when the information needed to be retrieved: approximately 4s when the needed database was decluttered relative to the fixed monochrome display and 2.5s relative to the fixed highlighting display when participants were given the ability to interactively highlight different information domains.

The data reporting the frequency of use of the adaptive and interactive display functions (in Table 4) suggests that the utility of the function was primarily to retrieve information, when it was not already present, rather than to customize the display. These results are similar to those reported by Oviatt (1996). This data, taken into account with the data in Table 5, suggest that the presence of the interactivity tools prolong response time, with no effect on accuracy. What is interesting is the fact that participants used the interactivity tools at all, given that it did not appear to help them. However since the interaction was discretionary for several question categories, we cannot establish if, for those subjects who used it, their response times might have

been longer (or their accuracy higher) on those particular questions, had this interaction been prevented.

GENERAL DISCUSSION

Fixed Displays

The current data reveal that some aspects of serial visual search models appear to be reasonably accurate in predicting response time performance, at least at a global scale. However the precise mechanism of the search operation remains unclear, and is the subject of future research. The current experiment informs different theories of attentional filtering. In general, the augmented fixed displays supported better performance than the monochrome display. First, the accuracy of responses was significantly improved using the color display relative to the monochrome displays (particularly in both Experiments 1 and 2). Correspondingly, intensity coding resulted in performance that was as good as that obtained with the color-coded display. Although there was no statistically significant difference between the intensity coding and monochrome displays in Experiment 1, there was a significant benefit for one of the highlighting conditions (people highlighted) as was shown in Experiment 2. Furthermore, both intensity coding conditions buffered the effect of clutter on the accuracy of answering divided attention conditions, a penalty that was observed in the monochrome condition.

Surprisingly, the benefits for the two techniques were not stronger, given the large benefits observed for intensity coding by Martens and Wickens (1995). In explaining the reduced benefits observed here, it should be noted that the two sets of features (linear vs. people) were fairly discriminable from each other by salient spatial characteristics. Roads were thick dotted lines and rivers thick solid lines, whereas troops were represented by arrows and units by shapes (squares and triangles). Furthermore, it should be noted that the two domains and classes within each domain were created to be semantically relevant and cognitively distinct. Linear features were environmental elements consisting of rivers, which could be viewed a “natural” occurrence, and roads, “man-made” objects. People features were Army personnel who were part of units, stationary checkpoints, or troops, traveling soldiers. Thus, this distinction might have reinforced the physical differences between the two classes, even in the absence of coding.

The data revealed that linear features were easier to search for and locate than people features, a difference that can accent for the selective response time benefit of highlighting people features, but not linear features in Experiment 2. This advantage for linear features was unaffected by whether the features were highlighted or uniquely color coded. There are two possible interpretations for this “feature effect.” First, and perhaps the most straight-forward explanation, it was the case that most maps generally contained more people featured items than linear featured items. If searching within a coded class (linear or people) proceeded in a serial manner, then one would predict, on average, longer search times for focused attention and integrated questions pertaining to people than to linear features. Second, the results may imply that rather than confusing visual features, subjects seemed to confuse “mental” features of items within a domain. The grouping of “people” features may have forced subjects to make a distinction between two unfamiliar entities, troop movement and units. The key fact to remember was that troops moved whereas units were stationary, and participants needed to apply this information to draw conclusions about the two classes of information. Thus, this subtle

distinction may have resulted in increasing the response time whereas the distinction between the two classes of linear features -- water and roads -- was made more intuitively as the two classes were more familiar to subjects.

While the creation of feature differences through color and intensity coding did not greatly facilitate search for those features in focused and divided attention questions, it is important to note that neither did these augmentation techniques disrupt the ability to divide attention between them to answer the integration questions, in contrast to results found in experiments performed by Andre and Wickens (1989). However, the results of the current experiment were not as strong as those found by Martens and Wickens (1995) who actually showed an advantage for highlighting with divided attention tasks such that response time to questions decreased as the intensity differences between the domains increased. Martens and Wickens hypothesized that the divided attention task was actually comprised of two focused attention search tasks, whose performance was facilitated by the intensity differences between the domains, followed by an integration task. Thus, the differing intensities actually helped to partition the two domains. The results of the current experiment support this hypothesis: the divided attention tasks required participants to first focus on objects within two to three different domains, which were discriminable in color or intensity. Once the objects were found, they could easily be compared and integrated. Hence, we infer that the advantages for coding different display elements distinctly, when these elements need to be compared or integrated, will grow as the search demands of the task increase. This inference is consistent with the difference between the results on the one hand of Andre and Wickens (1989) and those of Martens and Wickens (1995) and the current data on the other. It is also consistent with the accuracy data observed in Experiment 2 (Figure 4), in which coding by either intensity or color reduced or eliminated the accuracy costs on divided attention questions, which had been observed in the monochrome display.

Interactive and Adaptive Displays

Changing the focus of our discussion to the interactive displays, we examine the results in which decluttering was contrasted with interactive intensity coding. These two techniques were grouped together as both allowed participants to “de-emphasize” one set of features relative to the other. In the case of intensity coding, this was accomplished by dimming one domain; with decluttering, the domain was erased altogether. Additionally, the use of intensity coding allowed an attentional filter to parse the information space whereas with decluttering, manual intervention was required.

The data supported the conclusion that there was an advantage for “dimming” over interactive “decluttering”. A comparison of the interactive decluttering (Table 5a) vs. fixed highlighted data (Table 5b) show no advantage for one technique over the other when the needed information was present, i.e., there was no cost for attentional filtering of lowlighted data, but a significant cost when the information needed to be retrieved. For divided attention questions, there was an advantage for having information present in a constant format and even a slight benefit for the fact that it is at a different intensity. The data emphasizes the cost of information retrieval, such that the time spent initially making a decision as to whether the necessary information is present and the retrieval of the necessary information adds a penalty to the use of either decluttering or interactive highlighting with no apparent gain in accuracy.

Additionally, we note one source of confusion in decluttering may be the manual responses required to display the initially decluttered (erased) information. In our study, participants were asked to make only two inputs: (1) pressing the space bar to answer a question or (2) press a key to display the linear features or to display the people features. We observed instances in which the subject would hit the space bar accidentally in an attempt to bring up "decluttered" information due to the similarity of the responses (pressing a key). It is important to note that this phenomenon could be amplified under stressful conditions, e.g., the task of the pilot who attempts to declutter his approach chart during the high workload periods of landing.

CONCLUSION

Results from experiments conducted in controlled laboratory environment were examined to determine how well they predicted search in an applied setting. Manipulations in the display design were made in an effort to determine how to best present tactical information in cluttered environments. The application of these results is not limited to cartographers but to user interface artists and all those who strive to improve the design of cluttered displays.

The main effect of display for both static maps and flexible maps emphasize the value of segregating the items in the visual field to facilitate search tasks using intensity coding and particularly color coding, for example. The use of interactive decluttering or dimming techniques to reduce the clutter in the visual field seemed to "cause more harm than good". The necessary retrieval of information not only added one to three seconds to the response time but also created confusion in the selection of the manual response. This limits the application of decluttering to conditions in which the time critical responses are not necessary. Even the flexibility to selectively highlight or dim classes of information appears to impose a time cost (of deciding whether such an option is needed) with no corresponding benefit to accuracy. Attention filtering can often do a more effective job than interactive filtering, in bringing the useful information to mind at the appropriate time.

Future research to be performed will further investigate the effect of decluttering on dynamic displays, particularly if the domain that is "hidden" contains information that changes (e.g., the appearance of an aircraft) to the user's task. Additionally, we attempt to examine the efficiency of these techniques with higher degrees of clutter. By varying the amount of information on the displays, we hope to better quantify the meaning of clutter.

ACKNOWLEDGEMENTS

The authors wish to thank Martha Weller for her programming assistance and David Brandenburg, James Orr, and Jennifer Stephens for their assistance in the data collection. Research funding for this project was provided by a grant from the U.S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0003. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Army Research Laboratory.

REFERENCES

Andre, A.D. and Wickens, C.D. (1989). *Information processing and perceptual characteristics of display design: The role of emergent features and objects* (Technical Report ARL-89-8/AHEL-89-4). Savoy, IL: Aviation Research Laboratory.

Fisher, D.L., Coury, B.G., Tengs, T.O., and Duffy, S.A. (1989). Minimizing the time to search visual displays: The role of highlighting. *Human Factors, 31*(2), 167-182.

Fisher, D.L. and Tan, K.C. (1989). Visual displays: The highlighting paradox. *Human Factors*, *31*(1), 17-30.

Hofer, E.F., Palen, L.A., and Possolo, A. (1993). Flight deck information management: An experimental study of functional integration of approach data. *Proceedings of the 7th International Symposium on Aviation Psychology*. Columbus, OH: Department of Aviation, Ohio State University.

Martens, M.H. and Wickens, C.D. (1995). Intensity coding solutions to display clutter (Technical Report ARL 95-10/NASA 95-4). Savoy, IL: Aviation Research Laboratory.

Mykityshyn, M.G., Kuchar, J.K, and Hansman, R.J. (1994). Experimental study of electronically based instrument approach plates. *The International Journal of Aviation Psychology*. *4*(2), 141-166.

Neisser, U., Novick, R., and Lazar, R. (1964). Searching for novel targets. *Perceptual and Motor Skills*, *19*, 427-432.

Oviatt, S. (1996). Multimodal interfaces for dynamic interactive maps. *Proceedings of Chi '96* (pp. 95-102). April 13-18, 1996: Vancouver, BC Canada.

Schons, V. and Wickens, C.D. (1993). Visual separation and information access in aircraft display layout (Technical Report ARL-93-7/NASA-A^3I-93-1). Savoy, IL: Aviation Research Laboratory.

Schultz, E.E. (1986). Decluttering methods for computer generated graphic displays. *NASA Technical Brief 10(3), Item #40*. Jet Propulsion Laboratory: Pasadena, CA.

Schultz, E.E., Nickols, D.A., and Curran, P.S. (1985). Decluttering methods for high-density computer-generated graphic displays. *Proceedings of the 29th Annual Meeting of the Human Factors Society*. Santa Monica, CA: Human Factors Society.

Teichner, W.H. and Mocharnuk, J.B. (1979). Visual search for complex targets. *Human Factors*. *21*(3), 259-275.

Treisman, A. (1988). Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology*. *40A*(2), 201-237.

Treisman, A.M., Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136.

Tullis, T.S. (1988). Screen design. In M. Helander (ed.), *Handbook of human-computer interaction* (pp. 377-411). Amsterdam: North-Holland.

Vincow, M.A. and Wickens, C.D. (1993). Spatial layout of displayed information: Three steps toward developing quantitative models. *Proceedings of the 37th Annual Meeting of the Human Factors Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Wickens, C.D. and Andre, A.D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness of information integration. *Human Factors, 32*, 61-77.

Wickens, C. and Carswell, C.M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, *37*(3), 473-494.

Wickens, C.D. and Hollands, J.G. (2000). *Engineering psychology and human performance*. Upper Saddle River, NJ: Prentice Hall.

Wickens, C.D., Vincow, M.A., Schopper, A.W., and Lincoln, J.E. (1997). *Computational Models of Human Performance in the Design and Layout of Controls and Displays*. Wright-Patterson AFB, OH: CSERIAC.

Wickens, C.D. and Yeh, M. (1997). Attentional filtering and decluttering techniques in battlefield map interpretation. *Proceedings of the Advanced Displays and Interactive Displays Federated Laboratory First Annual Symposium* (2-35 - 2-42). Adelphi, MD: Army Research Laboratory.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2*, 156-161.

Yeh, M., & Wickens, C.D. (1997). *Performance issues in helmet-mounted displays* (Technical Report ARL-97-9/ARMY-FED-LAB-97-1). Savoy, IL: Aviation Res. Lab.